

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- 5 (i) APPLICANT: Washington University
- (ii) TITLE OF INVENTION: BLNK PROTEINS
- 10 (iii) NUMBER OF SEQUENCES: 12
- (iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
- 15 (B) STREET: Four Embarcadero Center, Suite 3400
- (C) CITY: San Francisco
- (D) STATE: California
- (E) COUNTRY: United States
- (F) ZIP: 94111-4187
- (v) COMPUTER READABLE FORM:
- 20 (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- 25 (vi) CURRENT APPLICATION DATA:
- (A) APPLICATION NUMBER: PCT/US98/
- (B) FILING DATE: Herewith
- (C) CLASSIFICATION:
- 30 (vii) PRIOR APPLICATION DATA:
- (A) APPLICATION NUMBER: US 08/819,013
- (B) FILING DATE: 17-MAR-1997
- (vii) PRIOR APPLICATION DATA:
- 35 (A) APPLICATION NUMBER: US 08/788,322
- (B) FILING DATE: 24-JAN-1997
- (viii) ATTORNEY/AGENT INFORMATION:
- 40 (A) NAME: Silva, Robin M.
- (B) REGISTRATION NUMBER: 38,304
- (C) REFERENCE/DOCKET NUMBER: FP-64383-2/RFT/RMS
- (ix) TELECOMMUNICATION INFORMATION:
- 45 (A) TELEPHONE: (415) 781-1989
- (B) TELEFAX: (415) 398-3249
- (C) TELEX: 910 277299
- (2) INFORMATION FOR SEQ ID NO:1:
- 50 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 456 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- 55 (D) TOPOLOGY: unknown

-29-

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

5 Met Asp Lys Leu Asn Lys Ile Thr Val Pro Ala Ser Gln Lys Leu Arg
 1 5 10 15
 Gln Leu Gln Lys Met Val His Asp Ile Lys Asn Asn Glu Gly Gly Ile
 20 25 30
 10 Met Asn Lys Ile Lys Lys Leu Lys Val Lys Ala Pro Pro Ser Val Pro
 35 40 45
 15 Arg Arg Asp Tyr Ala Ser Glu Ser Pro Ala Asp Glu Glu Glu Gln Trp
 50 55 60
 Ser Asp Asp Phe Asp Ser Asp Tyr Glu Asn Pro Asp Glu His Ser Asp
 65 70 75 80
 20 Ser Glu Met Tyr Val Met Pro Ala Glu Glu Asn Ala Asp Asp Ser Tyr
 85 90 95
 Glu Pro Pro Pro Val Glu Gln Glu Thr Arg Pro Val His Pro Ala Leu
 100 105 110
 25 Pro Phe Ala Arg Gly Glu Tyr Ile Asp Asn Arg Ser Ser Gln Arg His
 115 120 125
 Ser Pro Pro Phe Ser Lys Thr Leu Pro Ser Lys Pro Ser Trp Pro Ser
 130 135 140
 Glu Lys Ala Arg Leu Thr Ser Thr Leu Pro Ala Leu Thr Ala Leu Gln
 145 150 155 160
 35 Lys Pro Gln Val Pro Pro Lys Pro Lys Gly Leu Leu Glu Asp Glu Ala
 165 170 175
 Asp Tyr Val Val Pro Val Glu Asp Asn Asp Glu Asn Tyr Ile His Pro
 180 185 190
 40 Thr Glu Ser Ser Ser Pro Pro Pro Glu Lys Ala Pro Met Val Asn Arg
 195 200 205
 Ser Thr Lys Pro Asn Ser Ser Thr Pro Ala Ser Pro Pro Gly Thr Ala
 210 215 220
 Ser Gly Arg Asn Ser Gly Ala Trp Glu Thr Lys Ser Pro Pro Pro Ala
 225 230 235 240
 50 Ala Pro Ser Pro Leu Pro Arg Ala Gly Lys Lys Pro Thr Thr Pro Leu
 245 250 255
 Lys Thr Thr Pro Val Ala Ser Gln Gln Asn Ala Ser Ser Val Cys Glu
 260 265 270
 55

652224-000000

-30-

Glu Lys Pro Ile Pro Ala Glu Arg His Arg Gly Ser Ser His Arg Gln
 275 280 285
 5 Glu Ala Val Gln Ser Pro Val Phe Pro Pro Ala Gln Lys Gln Ile His
 290 295 300
 Gln Lys Pro Ile Pro Leu Pro Arg Phe Thr Glu Gly Gly Asn Pro Thr
 305 310 315 320
 10 Val Asp Gly Pro Leu Pro Ile Phe Ser Ser Asn Ser Thr Ile Ser Glu
 325 330 335
 Gln Glu Ala Gly Val Leu Cys Lys Pro Trp Tyr Ala Gly Ala Cys Asp
 340 345 350
 15 Arg Lys Ser Ala Glu Glu Ala Leu His Arg Ser Asn Lys Asp Gly Ser
 355 360 365
 Phe Leu Ile Arg Lys Ser Ser Gly His Asp Ser Lys Gln Pro Tyr Thr
 370 375 380
 Leu Val Val Phe Phe Asn Lys Arg Val Tyr Asn Ile Pro Val Arg Phe
 385 390 395 400
 25 Ile Glu Ala Thr Lys Gln Tyr Ala Leu Gly Arg Lys Lys Asn Gly Glu
 405 410 415
 Glu Tyr Phe Gly Ser Val Ala Glu Ile Ile Arg Asn His Gln His Ser
 420 425 430
 30 Pro Leu Val Leu Ile Asp Ser Gln Asn Asn Thr Lys Asp Ser Thr Arg
 435 440 445
 Leu Lys Tyr Ala Val Lys Val Ser
 450 455
 35

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- 40 (A) LENGTH: 1806 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

45 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

50 CCTTCGTGGC CGCAGCCTGC ACTCTCAGAA ATCAGACTTG AGTGGCCGGA ACCCTTGAGA 60
 CCAGAGGCTT ACCATGCTGC TCCCTAGGAG GGCCAGGAAC TGCTGACGTG ACCACTGGAC 120
 AGTTATTCGT GTCTCTTACA ATTACCAAAC AGAATGGACA AGCTTAATAA AATAACCGTC 180
 55 CCCGCCAGTC AGAAGTTGAG GCAGCTTCAA AAGATGGTCC ATGATATTAA AAACAATGAA 240

0935440330
 6624042550

GGTGGAATAA TGAATAAAAT CAAAAAGCTA AAAGTCAAAG CACCTCCAAG TGTTCCTCGA 300
AGGGACTACG CTTTCAGAGAG CCCCCTGAC GAAGAGGAGC AGTGGTCCGA TGAATTTGAC 360
5 AGCGACTATG AAAATCCAGA TGAGCACTCG GACTCAGAGA TGTACGTGAT GCCCGCCGAG 420
GAGAACGCTG ATGACAGCTA CGAGCCGCCT CCAGTAGAGC AGGAAACCAG GCCGGTTCAC 480
10 CCAGCCCTGC CCTTCGCCAG AGGCGAGTAT ATAGACAATC GATCAAGCCA GAGGCATTCC 540
CCACCCTTCA GCAAGACACT TCCCAGTAAG CCCAGCTGGC CTTTCAGAGAA AGCAAGGCTC 600
ACCTCCACCC TGCCGGCCCT GACTGCTTTG CAGAAACCTC AAGTCCCACC CAAACCCAAA 660
15 GGCCTCCTTG AGGATGAGGC TGATTATGTG GTCCCCGTGG AAGATAATGA TGAAACTAT 720
ATTCATCCCA CAGAAAGCAG TTCACCTCCA CCTGAAAAAG CTCCCATGGT GAATAGATCA 780
ACCAAGCCAA ATTCCTCAAC GCCCGCCTCT CCTCCAGGAA CAGCTTCAGG TCGAAACAGT 840
20 GGGGCCTGGG AAACCAAGTC ACCTCCACCA GCTGCACCAT CCCCCTTGCC ACGGGCCGGG 900
AAAAAACCAA CGACACCACT GAAGACAACT CCAGTTGCCT CTCAACAGAA TGCTTCAAGT 960
25 GTTTGTGAAG AAAAACCTAT ACCTGCTGAA CGCCACCGAG GGTCAAGTCA CAGACAAGAA 1020
GCTGTGCAGT CACCAGTGTT TCCTCCTGCC CAGAAACAAA TCCACCAAAA ACCCATACTT 1080
CTGCCAAGAT TTACAGAAGG GGGAAACCCA ACTGTGGATG GGCCCCTACC CAGCTTTTCA 1140
30 TCTAATTCCA CTATTTTCTA ACAGGAAGCT GGCCTTCTCT GCAAGCCATG GTATGCTGGA 1200
GCCTGTGATC GAAAGTCTGC TGAAGAGGCA TTGCACAGAT CAAACAAGGA TGGATCATTT 1260
35 CTTATTCGGA AAAGCTCTGG CCATGATTCC AAACAACCAT ATACACTAGT TGTATTCTTT 1320
AATAAGCGAG TATATAATAT TCCTGTGCGA TTTATTGAAG CAACAAAACA ATATGCCTTG 1380
GGCAGAAAGA AAAATGGTGA AGAGTACTTT GGAAGTGTTG CTGAAATCAT CAGGAATCAT 1440
40 CAACATAGTC CTTTGGTTCT TATTGACAGT CAGAATAACA CAAAAGATTC CACCAGACTG 1500
AAGTATGCAG TTAAAGTTTC ATAAAGGGGG AAAAAAAGA TCAATACCAT TGCTTCAGAC 1560
45 ACTTTCCCAA AGTTTCTCCT TTTGAGAAAA AGTCCCAAAA CTTTCATATT TGGATTATGA 1620
ATCATCCAGT AATAAAATGG AAGATGGAGT CAGCTATTGA AGTGGTCATC CATTTCTTTT 1680
TAAGAAGCTC ATGTGGACTT GTTCTATTGC CTGACCTGAT GAACTGTTAA TATCTGGTGA 1740
50 GGTGAGTTA TCATGCTACT AATATTTTCC AAATAAATAT TTTTATTTTT AAAAAAAAAA 1800
AAAAAA 1806

55 (2) INFORMATION FOR SEQ ID NO:3:

-32-

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 449 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Asp Lys Leu Asn Lys Ile Thr Val Pro Ala Ser Gln Lys Leu Arg
 1 5 10 15

His Ile Lys Asn Asn Glu Gly Gly Ile Met Asn Lys Ile Lys Lys Leu
 20 25 30

Lys Val Lys Ala Pro Pro Ser Val Pro Arg Arg Asp Tyr Ala Ser Glu
 35 40 45

Ser Pro Ala Asp Glu Glu Glu Gln Trp Ser Asp Asp Phe Asp Ser Asp
 50 55 60

Tyr Glu Asn Pro Asp Glu His Ser Asp Ser Glu Met Tyr Val Met Pro
 65 70 75 80

Ala Glu Glu Asn Ala Asp Asp Ser Tyr Glu Pro Pro Pro Val Glu Gln
 85 90 95

Glu Thr Arg Pro Val His Pro Ala Leu Pro Phe Ala Arg Gly Glu Tyr
 100 105 110

Ile Asp Asn Arg Ser Ser Gln Arg His Ser Pro Pro Phe Ser Lys Thr
 115 120 125

Leu Pro Ser Lys Pro Ser Trp Pro Ser Glu Lys Ala Arg Leu Thr Ser
 130 135 140

Thr Leu Pro Ala Leu Thr Ala Leu Gln Lys Pro Gln Val Pro Pro Lys
 145 150 155 160

Pro Lys Gly Leu Leu Glu Asp Glu Ala Asp Tyr Val Val Pro Val Glu
 165 170 175

Asp Asn Asp Glu Asn Tyr Ile His Pro Thr Glu Ser Ser Ser Pro Pro
 180 185 190

Pro Glu Lys Ala Pro Met Val Asn Arg Ser Thr Lys Pro Asn Ser Ser
 195 200 205

Thr Pro Ala Ser Pro Pro Gly Thr Ala Ser Gly Arg Asn Ser Gly Ala
 210 215 220

Trp Glu Thr Lys Ser Pro Pro Pro Ala Ala Pro Ser Pro Leu Pro Arg
 225 230 235 240

656220-4729260

-33-

Ala Gly Lys Lys Pro Thr Thr Pro Leu Lys Thr Thr Pro Val Ala Ser
245 250 255

5 Gln Gln Asn Ala Ser Ser Val Cys Glu Glu Lys Pro Ile Pro Ala Glu
260 265 270

Arg His Arg Gly Ser Ser His Arg Gln Glu Ala Val Gln Ser Pro Val
275 280 285

10 Phe Pro Pro Ala Gln Lys Gln Ile His Gln Lys Pro Ile Pro Leu Pro
290 295 300

Arg Phe Thr Glu Gly Gly Asn Pro Thr Val Asp Gly Pro Leu Pro Ser
305 310 315 320

15 Phe Ser Ser Asn Ser Thr Ile Ser Glu Gln Glu Ala Gly Val Leu Cys
325 330 335

Lys Pro Trp Tyr Ala Gly Ala Cys Asp Arg Lys Ser Ala Glu Glu Ala
20 340 345 350

Leu His Arg Ser Asn Lys Asp Gly Ser Phe Leu Ile Arg Lys Ser Ser
355 360 365

25 Gly His Asp Ser Lys Gln Pro Tyr Thr Leu Val Val Phe Phe Asn Lys
370 375 380

Arg Val Tyr Asn Ile Pro Val Arg Phe Ile Glu Ala Thr Lys Gln Tyr
385 390 395 400

30 Ala Leu Gly Arg Lys Lys Asn Gly Glu Glu Tyr Phe Gly Ser Val Ala
405 410 415

Glu Ile Ile Arg Asn His Gln His Ser Pro Leu Val Leu Ile Asp Ser
35 420 425 430

Gln Asn Asn Thr Lys Asp Ser Thr Arg Leu Lys Tyr Ala Val Lys Val
435 440 445

40 Ser

(2) INFORMATION FOR SEQ ID NO:4:

45 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1785 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

50 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

55 CCTTCGTGGC CGCAGCCTGC ACTCTCAGAA ATCAGACTTG AGTGGCCGGA ACCCTTGAGA

60

CCAGAGGCTT ACCATGCTGC TCCCTAGGAG GGCCAGGAAC TGCTGACGTG ACCACTGGAC 120
AGTTATTCGT GTCTCTTACA ATTACCAAAC AGAATGGACA AGCTTAATAA AATAACCGTC 180
5 CCCGCCAGTC AGAAGTTGAG GCATATTAAA AACAATGAAG GTGGAATAAT GAATAAAATC 240
AAAAAGCTAA AAGTCAAAGC ACCTCCAAGT GTTCCTCGAA GGGACTACGC TTCAGAGAGC 300
CCCGCTGACG AAGAGGAGCA GTGGTCCGAT GACTTTGACA GCGACTATGA AAATCCAGAT 360
10 GAGCACTCGG ACTCAGAGAT GTACGTGATG CCCGCCGAGG AGAACGCTGA TGACAGCTAC 420
GAGCCGCCTC CAGTAGAGCA GGAAACCAGG CCGGTTCCACC CAGCCCTGCC CTTCCGCCAGA 480
15 GGCGAGTATA TAGACAATCG ATCAAGCCAG AGGCATTCCC CACCCTTCAG CAAGACACTT 540
CCCAGTAAGC CCAGCTGGCC TTCAGAGAAA GCAAGGCTCA CCTCCACCCT GCCGGCCCTG 600
ACTGCTTTGC AGAAACCTCA AGTCCCACCC AAACCCAAAG GCCTCCTTGA GGATGAGGCT 660
20 GATTATGTGG TCCCCGTGGA AGATAATGAT GAAAACTATA TTCATCCCAC AGAAAGCAGT 720
TCACCTCCAC CTGAAAAAGC TCCCATGGTG AATAGATCAA CCAAGCCAAA TTCCTCAACG 780
25 CCCGCCTCTC CTCCAGGAAC AGCTTCAGGT CGAAACAGTG GGGCCTGGGA AACCAAGTCA 840
CCTCCACCAG CTGCACCATC CCCGTTGCCA CGGGCCGGGA AAAAACCAAC GACACCACTG 900
AAGACAACTC CAGTTGCCTC TCAACAGAAT GCTTCAAGTG TTTGTGAAGA AAAACCTATA 960
30 CCTGCTGAAC GCCACCGAGG GTCAAGTCAC AGACAAGAAG CTGTGCAGTC ACCAGTGTTT 1020
CCTCCTGCCC AGAAACAAAT CCACCAAAAA CCCATACCTC TGCCAAGATT TACAGAAGGG 1080
35 GGAAACCCAA CTGTGGATGG GCCCCTACCC AGCTTTTCAT CTAATTCCAC TATTTTCAGAA 1140
CAGGAAGCTG GCGTTCTCTG CAAGCCATGG TATGCTGGAG CCTGTGATCG AAAGTCTGCT 1200
GAAGAGGCAT TGCACAGATC AAACAAGGAT GGATCATTTT TTATTCGGAA AAGCTCTGGC 1260
40 CATGATTCCA AACAACCATA TACACTAGTT GTATTCTTTA ATAAGCGAGT ATATAATATT 1320
CCTGTGCGAT TTATTGAAGC AACAAAACAA TATGCCTTGG GCAGAAAGAA AAATGGTGAA 1380
45 GAGTACTTTG GAAGTGTTGC TGAAATCATC AGGAATCATC AACATAGTCC TTTGGTTCTT 1440
ATTGACAGTC AGAATAACAC AAAAGATTCC ACCAGACTGA AGTATGCAGT TAAAGTTTCA 1500
TAAAGGGGGA AAAAAAAGAT CAATACCATT GCTTCAGACA CTTTCCCAA GTTTCTCCTT 1560
50 TTGAGAAAAA GTCCCAAAC TTCATATTTT GGATTATGAA TCATCCAGTA ATAAAATGGA 1620
AGATGGAGTC AGCTATTGAA GTGGTCATCC ATTTCTTTTT AAGAAGCTCA TGTGGACTTG 1680
55 TTCTATTGCC TGACCTGATG AACTGTTAAT ATCTGGTGAG GTTGAGTTAT CATGCTACTA 1740

6620445560

1785

5 (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 457 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

15	Met	Asp	Lys	Leu	Asn	Lys	Ile	Thr	Val	Pro	Ala	Ser	Gln	Lys	Leu	Arg
	1				5					10					15	
	Gln	Leu	Gln	Lys	Met	Val	His	Asp	Ile	Lys	Asn	Asn	Glu	Gly	Gly	Ile
				20					25					30		
20	Met	Asp	Lys	Ile	Lys	Lys	Leu	Lys	Val	Lys	Gly	Pro	Pro	Ser	Val	Pro
			35					40					45			
	Arg	Arg	Asp	Tyr	Ala	Leu	Asp	Ser	Pro	Ala	Asp	Glu	Glu	Glu	Gln	Trp
25		50					55					60				
	Ser	Asp	Asp	Phe	Asp	Ser	Asp	Tyr	Glu	Asn	Pro	Asp	Glu	His	Ser	Asp
	65					70					75					80
30	Ser	Glu	Met	Tyr	Val	Met	Pro	Ala	Glu	Glu	Thr	Gly	Asp	Asp	Ser	Tyr
					85					90					95	
	Glu	Pro	Pro	Pro	Ala	Glu	Gln	Gln	Thr	Arg	Val	Val	His	Pro	Ala	Leu
				100					105					110		
35	Pro	Phe	Thr	Arg	Gly	Glu	Tyr	Val	Asp	Asn	Arg	Ser	Ser	Gln	Arg	His
			115					120					125			
	Ser	Pro	Pro	Phe	Ser	Lys	Thr	Leu	Pro	Ser	Lys	Pro	Ser	Trp	Pro	Ser
40		130					135					140				
	Ala	Lys	Ala	Arg	Leu	Ala	Ser	Thr	Leu	Pro	Ala	Pro	Asn	Ser	Leu	Gln
	145					150					155					160
45	Lys	Pro	Gln	Val	Pro	Pro	Lys	Pro	Lys	Asp	Leu	Leu	Glu	Asp	Glu	Ala
					165					170					175	
	Asp	Tyr	Val	Val	Pro	Val	Glu	Asp	Asn	Asp	Glu	Asn	Tyr	Ile	His	Pro
				180					185					190		
50	Arg	Glu	Ser	Ser	Pro	Pro	Pro	Ala	Glu	Lys	Ala	Pro	Met	Val	Asn	Arg
			195					200					205			
	Ser	Thr	Lys	Pro	Asn	Ser	Ser	Ser	Lys	His	Met	Ser	Pro	Pro	Gly	Thr
55		210						215				220				

-36-

Val Ala Gly Arg Asn Ser Gly Val Trp Asp Ser Lys Ser Ser Leu Pro
 225 230 235 240
 5 Ala Ala Pro Ser Pro Leu Pro Arg Ala Gly Lys Lys Pro Ala Thr Pro
 245 250 255
 Leu Lys Thr Thr Pro Val Pro Pro Leu Pro Asn Ala Ser Asn Val Cys
 260 265 270
 10 Glu Glu Lys Pro Val Pro Ala Glu Arg His Arg Gly Ser Ser His Arg
 275 280 285
 Gln Asp Thr Val Gln Ser Pro Val Phe Pro Pro Thr Gln Lys Pro Val
 290 295 300
 15 His Gln Lys Pro Val Pro Leu Pro Arg Phe Pro Glu Ala Gly Ser Pro
 305 310 315 320
 Ala Ala Asp Gly Pro Phe His Ser Phe Pro Phe Asn Leu Thr Phe Ala
 325 330 335
 Asp Gln Glu Gly Glu Leu Leu Gly Lys Pro Trp Tyr Ala Gly Ala Cys
 340 345 350
 25 Asp Arg Lys Phe Ala Glu Glu Ala Leu His Arg Ser Asn Lys Asp Gly
 355 360 365
 Ser Phe Leu Ile Arg Lys Ser Phe Gly His Asp Ser Lys Gln Pro Tyr
 370 375 380
 30 Thr Leu Val Ala Phe Phe Asn Lys Arg Val Tyr Asn Ile Pro Val Arg
 385 390 395 400
 Phe Ile Glu Ala Thr Lys Gln Tyr Ala Leu Gly Lys Lys Lys Asn Gly
 405 410 415
 35 Glu Glu Tyr Phe Gly Ser Val Val Glu Ile Val Asn Ser His Gln His
 420 425 430
 40 Asn Pro Leu Val Leu Ile Asp Ser Gln Asn Asn Thr Lys Asp Ser Thr
 435 440 445
 Arg Leu Lys Tyr Ala Val Lys Val Ser
 450 455

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1718 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CTGTGGGTTG CTCGCAGAAG TCAGTTCCAG TGGCTTGAGT TCTTGAGGCC AGAGCCTTAC 60
CATGCTGCTC CCCAGGAAGT CCAGGAGCTG CTGACACCCC CCTGGACAGC GACACATCCT 120
5 CTCTCAAGAA AATGGACAAG CTGAATAAGA TAACTGTCCC TGCCAGCCAG AAGCTGAGAC 180
AGCTTCAAAA GATGGTCCAT GATATTAAGA ACAATGAAGG TGAATAATG GACAAGATAA 240
AAAAGCTAAA AGTCAAAGGC CCTCCAAGTG TTCCTCGAAG GGACTATGCA TTAGACAGCC 300
10 CTGCAGATGA AGAGGAGCAG TGGTCAGATG ACTTCGACAG TGACTATGAA AATCCAGATG 360
AACATTCGGA CTCCGAGATG TATGTGATGC CTGCCGAGGA GACGGGCGAC GATTCCTATG 420
15 AACCGCCTCC CGCTGAGCAG CAGACACGGG TGGTCCATCC AGCCCTGCCC TTCACGAGGG 480
GCGAGTATGT AGATAATCGA TCCAGCCAGC GGCCTCTCC GCCCTTCAGC AAGACACTTC 540
CCAGTAAGCC CAGCTGGCCT TCAGCGAAAG CGAGGCTGGC CTCCACTCTG CCAGCCCCCA 600
20 ACTCTCTACA GAAGCCTCAA GTCCCCCCA AGCCCAAAGA CCTCCTTGAG GATGAGGCTG 660
ATTATGTGGT CCCTGTGGAA GATAACGATG AAAACTATAT CCATCCCAGA GAAAGTAGCC 720
25 CGCCGCCTGC TGAGAAGGCT CCCATGGTGA ATAGATCAAC CAAGCCAAAC AGTTCCTCAA 780
AGCACATGTC GCCTCCAGGG ACTGTCGCAG GTCGAAACAG TGGGGTCTGG GACTCCAAGT 840
CATCTTTGCC TGCCGCACCA TCCCCACTAC CACGGGCTGG GAAGAAGCCA GCTACACCAC 900
30 TTAAGACTAC TCCCGTTCCT CCCCTACCGA ATGCATCAAA TGTTTGTGAA GAAAAGCCTG 960
TTCCTGCTGA GCGCCACCGA GGGTCTAGTC ACAGACAAGA CACTGTACAG TCACCAGTGT 1020
35 TTCCTCCCAC CCAGAAACCT GTCCATCAAA AGCCTGTACC CTTGCCAAGG TTCCCAGAAG 1080
CGGGGAGCCC AGCTGCAGAT GGACCGTTCC ACAGCTTCCC ATTTAATTTG ACGTTTGCAG 1140
ACCAGGAGGG TGAAGTCTC GGTAAGCCCT GGTATGCTGG CGCCTGTGAC CGCAAGTTTG 1200
40 CTGAAGAGGC CTTGCACAGA TCCAACAAGG ATGGATCGTT TCTTATTCGG AAGAGCTTTG 1260
GCCATGATTC CAAGCAGCCG TACACCCTAG TTGCGTTCTT TAACAAGCGA GTGTATAATA 1320
45 TTCCTGTACG GTTTATTGAA GCAACCAAAC AGTATGCTTT GGGAAAGAAG AAAAATGGTG 1380
AAGAGTACTT CGGAAGTGTT GTGGAAATCG TCAACAGTCA TCAGCACAAC CCCCTGGTTC 1440
TTATTGACAG TCAGAATAAC ACGAAAGATT CCACGAGACT GAAATATGCT GTGAAGGTTT 1500
50 CATAACGATA CCACGGTTCC AGACATGTCC TCTGTTTCTT CTTTGTGAGAA AACATCATAT 1560
TCTGGCTATG ACTCCTCAGC AGTAAGAGAG AAAAGATGAA TGAAGCCACT GAGGCTTCGT 1620
55 GAATGAATGA ATCTACTCCT TCCTAGGGCG TTCACACGAG CTTTCTATC ACCTGACCTG 1680

ACGAAGTCAT AGCTGGGGAG GTTCGGTTAC TATGATAC

1718

(2) INFORMATION FOR SEQ ID NO:7:

- 5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

10

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

15 Tyr Glu Xaa Pro
1

(2) INFORMATION FOR SEQ ID NO:8:

- 20 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

25

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- 30 (A) NAME/KEY: Modified-site
(B) LOCATION: 1..2
(D) OTHER INFORMATION: /note= "The 'X' appearing at
positions 1, 8 and 15, represent either Aspartic Acid or
Glutamic Acid."

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Xaa Xaa Tyr Glu Xaa Pro Xaa Xaa Xaa Tyr Xaa Xaa Pro Xaa Xaa Xaa
1 5 10 15

40 Tyr Glu Pro Pro
20

(2) INFORMATION FOR SEQ ID NO:9:

- 45 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

50

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- 55 (A) NAME/KEY: misc_feature
(B) LOCATION: 19..20

(D) OTHER INFORMATION: /note= "The 'N' at position 19 represents Inosine."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

5

TCGAGAATTC AARAARCCNA CNACNCC

27

(2) INFORMATION FOR SEQ ID NO:10:

10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 8 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: unknown

(D) TOPOLOGY: unknown

15

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

20

Lys Lys Pro Thr Thr Pro Leu Lys

1

5

(2) INFORMATION FOR SEQ ID NO:11:

25

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 27 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: unknown

(D) TOPOLOGY: unknown

30

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

(A) NAME/KEY: misc_feature

35

(B) LOCATION: 13..14

(D) OTHER INFORMATION: /note= "The 'N' at position 13 represents Inosine."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

40

CTGAGGATCC TTNGTNGCYT CDATRAA

27

(2) INFORMATION FOR SEQ ID NO:12:

45

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 15 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: unknown

(D) TOPOLOGY: unknown

50

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

55

Lys Arg Val Tyr Asn Ile Pro Val Arg Phe Ile Glu Ala Thr Lys

1

5

10

15